Mitigating Procrastination Problems Through a System of Planning and Reminders?

Haya Goldblatt

Advisor: Dave Owens

Abstract
We experimentally test whether a system of planning and reminders can reduce present-bias over real-effort tasks. If an individual makes a plan for when they will complete a series of tasks and they receive a reminder of that plan, do they show less present-bias than individuals who make a plan but receive no reminder? Our results suggest that receiving the reminder has no effect on present-bias; individuals who received the reminder procrastinated tasks at similar levels to those who did not view a reminder.

Acknowledgements
I would like to thank Andy Janco for crafting the website used for this experiment and for humoring me when I requested changes, as well as Dave Owens for his feedback and advice throughout this process. I am also grateful to Giri Parameswaran and Gabe Sekeres for their assistance and generosity in sharing their funding and making this experiment possible. Finally, I would like to thank my family and friends for their support.
1) Introduction

We have all experienced a lack of self-discipline in some domain of our lives. We plan to save money, or follow a diet, or write a paper, but when an opportunity to begin executing that plan arrives, we often push off the unpleasant activity and tell ourselves that we will do it on a later date. If this process repeats over time, individuals can engage in patterns of behavior that not only policy makers, but they themselves believe are bad for them.

Standard economic theory struggles to explain this behavior. If an agent chooses, say, to eat junk food over a salad, then standard theory concludes that eating junk food must have been their optimal choice. The agent revealed that their preference was to eat the junk food because agents necessarily make decisions that are in their best interest.

However, economists and psychologists today generally agree that individuals frequently have preferences that are “present-biased.” That is, individuals overvalue their happiness/comfort today relative to their happiness/comfort in the future. Accounting for present-biased preferences allows economists to make sense of the phenomenon where individuals make choices that are pleasant in the short-run but suboptimal in the long-run.

In particular, present-biased preferences can explain procrastination problems. Completing a task that is good for an individual in the long-run but has an immediate cost, such as saving money or quitting smoking, is unappealing in the present, and completing the task in the future appears optimal. But when the future date arrives, the agent shows present-bias and procrastinates to another future period. If the agent repeatedly believes that their future self will complete the activity and underestimates the degree of the temptation to procrastinate in the future, then the pattern of procrastination may repeat indefinitely, to their own detriment. Many strategies have been proposed to mitigate procrastination problems in these and other contexts,
from programs that automatically deposit earnings into a savings account (Thaler & Benartzi, 2004) to intermediate deadlines on long-term projects (Ariely & Wertenbroch, 2002). These strategies take a tempting option, such as spending money or delaying writing a paper, and either remove the temptation or make the temptation more costly.

A mechanism that makes future good behavior easier or future bad behavior harder is called a commitment device. If an individual is aware that he is present-biased and therefore will be tempted to procrastinate, he can use a commitment device to encourage his future self to make the decision that aligns with his current preferences. Commitment devices can be classified as “hard” or “soft” (Bryan et al., 2010). A hard commitment device has a real economic penalty. For example, if an individual is tempted to procrastinate saving money, she can set up a system that automatically deposits funds into a savings account and arrange to pay a fee if she makes withdrawals before a specified date. In contrast, a soft commitment device has primarily psychological consequences. Putting funds in an account called “children’s education” is a soft commitment device because withdrawing money for a vacation would make the agent experience disappointment or guilt (Bryan et al., 2010).

When designing a commitment device, harsher economic or psychological penalties are more likely to induce one’s future self to comply with one’s current preferences. But if the individual ultimately succumbs to the temptation, then they experience the penalty and therefore a high penalty may deter individuals from taking up the commitment device. Increasing the penalty involves a trade-off between the effectiveness of the device and take-up rates (Bryan et al., 2010). The optimal commitment device uses the minimum penalty that induces individuals to avoid a tempting option.
This paper tests the effectiveness of a soft commitment device with a low penalty. We examine procrastination problems when an agent must complete a sequence of short tasks over time. Specifically, we ask: if an individual makes a plan for when they will complete a series of tasks and they receive a reminder of their plan, will they procrastinate less than those who make a plan but do not receive a reminder? The reminder is a soft commitment device with a small penalty. If a participant departs from their plan, they may experience a degree of guilt, but they do not face any additional consequences. If this commitment device is successful, then a system of planning and reminders could be implemented in schools, businesses, or other organizations to mitigate procrastination problems. Such a system could reduce the costs associated with procrastination—it could lower stress levels, lead to improved performance on tasks, and even lead to higher income.

We hypothesize that if an individual receives a reminder of her past self’s plan to complete tasks on a particular schedule, she will show less present-bias. That is, the reminder reduces procrastination. We test this hypothesis experimentally following the methods of Augenblick et al. (2015).¹ Participants are recruited for a three-day experiment. They have to complete a series of short tasks between Day 2 and Day 3, but they choose how many tasks to allocate to each day. On Day 1 of the experiment, participants decide how many tasks to allocate to Day 2 and how many to allocate to Day 3. Then on Day 2, participants allocate tasks between Day 2 and Day 3, again.² Because tasks allocated to Day 2 on Day 2 must be completed in the present, a present-biased agent will be tempted to procrastinate by allocating fewer tasks to Day

---

¹ Augenblick et al. (2015)’s methods are not followed perfectly due to several practical limitations in our experiment. See Section 6 for a discussion of the key areas in which we depart from their design.

² One allocation is selected at random to be the “allocation-that-counts,” and participants must complete the tasks according to the schedule they outlined in the allocation-that-counts.
2. The difference between the number of tasks allocated to Day 2 on Day 1 and the number allocated to Day 2 on Day 2 is used to identify the degree of an individual’s present-bias, or the strength of their preference to procrastinate.

To test the hypothesis that receiving a reminder about one’s plan reduces procrastination, half of the participants receive a reminder and the other half do not. On Day 2 of the experiment, before allocating tasks between Day 2 and Day 3, the treatment group receives a reminder of the allocation decision they made on Day 1. After they have seen their original decision, they make their second allocation of tasks between Day 2 and Day 3. We test our hypothesis by comparing the strength of the preference to procrastinate in the treatment group to the strength of the preference to procrastinate in the control group.

Contrary to our hypothesis, we do not find evidence that receiving a reminder of one’s original allocation reduces present-bias. Perhaps surprisingly, we also do not definitively conclude that individuals have present-biased preferences, even when restricting to participants who did not receive a reminder. Because we largely follow the methods of Augenblick et al. (2015) and that study did find significant evidence of present-bias, we consider which differences between the experimental designs could explain the difference between our findings.

While several explanations are plausible, the most substantial difference between the designs is that our experiment took place online using Amazon Mechanical Turk, while their experiment was in-person with undergraduate participants. Differences between the formats and subject pools in experiments might account for some of the differences between findings.

2) Literature Review

---

3 We follow the β-δ model of present-biased preferences, where all non-present periods are underweighted in the utility function (O’Donoghue & Rabin, 1999). We explore this model in more detail in Section 2.
For many years, the prevailing model of an agent’s utility over time was “exponential
discounting” (Samuelson, 1937). In that model, utility in each period is discounted based on its
distance from the present. Discounting is captured by the parameter $\delta$; when considering utility $k$
periods in the future, an agent discounts utility in that period by $\delta^k$. If $\delta < 1$, then periods closer
to the present are valued more than periods farther in the future. Because utility in one period is
always discounted by $\delta$ relative to the previous period, an agent’s choice of how much utility to
allocate to each future period remains optimal over time. That is, preferences are “time-
consistent.” These time-consistent preferences do not allow for behavior like procrastination,
where completing an action on a future date appears optimal in the present, but when the future
date arrives, the agent’s preferences change.

Samuelson’s exponential discounting model was designed for mathematical convenience
rather than for capturing human behavior, and he stressed that the assumption of time-
consistency should be revised (Frederick & Loewenstein, 2002). Despite these warnings,
exponential discounting came to be the standard economic model of decision-making over time.
Motivated by consumers’ spending and saving behavior, Strotz (1956) suggested that consumer
preferences should be modeled as present-biased, instead. He proposed the first model involving
a discounting function that increases the weight of utility in a future period as the future period
approaches the present.

The prevailing mathematical model of present-biased preferences, and the model we use,
is the “hyperbolic discounting” method developed after Strotz (1956) (O’Donoghue & Rabin
1999, Laibson 1997). In this model, an agent’s lifetime utility is determined by their utility in
each period, $u_t$, and two parameters, $\delta$ and $\beta$: 
\[ U = u_0 + \beta \sum_{t=1}^{n} \delta^t u_t \]

where the parameter \( \delta \) is Samuelson’s exponential discounting parameter. The main contribution of the hyperbolic discounting model is the parameter \( \beta \), which allows an agent to discount all non-present periods at an additional fixed rate. Because \( \beta \) discounts all non-present periods equally, it captures bias for the present period. If \( \beta = 1 \), then this model is identical to Samuelson’s model and agents are time-consistent. If \( \beta < 1 \), then utility in all non-present periods is weighted less than utility in the present, and the agent is present-biased.

In this theoretical framework, O’Donoghue & Rabin (1999) studied an agent’s decision-making if the agent is unaware of his present-bias (“naive”) or perfectly aware of his present-bias (“sophisticated”). Because a sophisticated agent is aware of his bias, he knows that he will be tempted to procrastinate unpleasant tasks which could lead to an even greater cost in the future, whereas a naive agent is oblivious to this risk. Additional works extend this model to allow the agent to be aware that he is present-biased but underestimate the degree of his bias (“partially sophisticated”) (O’Donoghue & Rabin, 2001) or require the agent to complete several costly tasks before a future benefit is realized (O’Donoghue & Rabin, 2008).

Because present-biased agents who are partially or fully sophisticated anticipate that they will be tempted to procrastinate, they may make use of a commitment device to reduce or remove that temptation, whereas naive agents do not believe they have a need for commitment. Several works study present-biased preferences and the demand for commitment in applied settings.

In the domain of savings, Ashraf et al. (2006) performed a randomized control trial in the Philippines and found that savings accounts with restrictions on withdrawals were effective at
increasing savings. Participants could choose to prohibit withdrawals until deposits reached a certain level or until a specified period of time had passed. Because the account eliminated the ability to spend one’s savings until the requirement was satisfied, it operated as a hard commitment device.

Thaler & Benartzi (2004)’s Save More Tomorrow™ program employed several strategies, including a soft commitment device, to increase retirement savings. Employees’ savings were increased automatically as their salary rose, and because there was no penalty to opting-out of the program other than psychological consequences, this program used a soft commitment device. It was highly successful at increasing savings.4 Ambec & Treich (2007) found that rotating savings and credit associations (ROSCAs) are best understood as commitment devices. Members of a group contribute funds to a ROSCA each month, and at the end of the month the pot is awarded to one member. This system makes a member’s savings illiquid until she receives the pot (a hard commitment), and the individual is under pressure from the other members to continue contributing to the ROSCA (a soft commitment).

Agents also demand commitment devices in the domain of health. Della Vigna & Malmendier (2006) found that individuals who bought unlimited memberships to health clubs paid more per visit than they would if they had purchased smaller packages of classes. While their preferred explanation for this phenomenon was that individuals were overconfident about their future willingness to go to the gym, an additional explanation is that those individuals purchased memberships to increase the psychological cost of staying home and commit themselves to go to the gym. Trope & Fishbach (2000) also found that individuals demand commitment if they must experience an immediate cost to realize a future health benefit.

---

4 Save More Tomorrow™ involved several features that are likely to contribute to its ability to raise savings, so the success cannot be fully attributed to the commitment device aspect of the plan.
Participants were told to follow a certain diet to learn the optimal glucose level for their cognitive performance, and when the diet was more unpleasant, participants generally imposed harsher penalties to reduce the temptation to break the diet.5

In academic or work settings in which individuals complete effortful tasks, deadlines have been used as hard commitment devices. Ariely & Wertenbroch (2002) experimentally tested the demand for commitment; they had participants proofread three papers over three weeks. When participants were allowed to set intermediate deadlines for themselves, they did set deadlines significantly earlier than the final deadline, so participants demanded this hard commitment. Additional subjects were randomly assigned to complete the tasks either with intermediate deadlines or with only the final deadline, and because those with intermediate deadlines showed stronger performance, the authors concluded that imposing deadlines as commitment devices is an effective way to improve performance on real effort tasks.

While Ariely & Wertenbroch (2002) favored intermediate deadlines as commitment devices, Bisin & Hyndman (2020) were less conclusive. They required participants to complete three tasks over two weeks, and like Ariely & Wertenbroch (2002), Bisin & Hyndman (2020) used tasks that were tedious—each task involved transcribing and alphabetizing approximately 150 words. They found that among participants who had the option of setting intermediate deadlines, those who chose to impose the commitment device completed fewer tasks than those who did not. This result contrasts with Ariely & Wertenbroch (2002)’s support for deadlines as commitment devices.

---

5 Individuals increased the penalty for failing to complete an unpleasant task (such as following a diet) as the unpleasantness of the task increased, but only up to a point. In separate experiments, Trope & Fishbach (2000) showed that if a task is too unpleasant, participants did not attempt to complete the task and did not use any self-control techniques.
Like Ariely & Wertenbroch (2002) and Bisin & Hyndman (2020), we study present-bias and commitment devices in a context where individuals must complete a series of real effort tasks. However, the commitment device we test is a soft commitment involving a system of planning and reminders, rather than a hard commitment involving deadlines.

Even without reminders, the formation of a plan can help individuals make decisions in the future that align with their current preferences. Gollwitzer & Sheeran (2006) conducted a meta-analysis on the effectiveness of “implementation intentions,” statements in the form “if Y occurs, then I will do X,” such as “if it is 9:00pm, then I will do my laundry.” Articulating a plan in this format increased the likelihood of accomplishing a variety of goals, from exercising more frequently to writing a CV. Implementation intentions could lead to these successes for a variety of reasons, and it is unclear whether a reduction in procrastination is one of those reasons. Valshtein et al. (2020) found that implementation intentions could reduce bedtime procrastination, the phenomenon of staying up late unnecessarily. But Gustavson & Miyake (2017) found no effect of implementation intentions on academic procrastination.

Separately from forming plans, reminders have been proposed as a solution to problems associated with present-bias. In a randomized field experiment, Karlan et al. (2010) found that receiving monthly reminders about a future expense significantly increased savings, and they argued that reminders increased savings because they increased the salience of the expense. An additional interpretation is that individuals planned to make monthly deposits, and receiving a reminder functioned as a commitment device that increased the psychological cost of spending.

Cadena et al. (2011) combined planning with reminders in an intervention to reduce procrastination among employees at a Colombian bank. Employees had some responsibilities that fell at the end of each month, and other jobs that could be completed at any time. Employees
generally procrastinated such that the second half of the month was much busier. Cadena et al. (2011) established a system in which the bank provided weekly targets for employees’ progress and offered small prizes for following that plan. The intervention significantly reduced procrastination, but only when the employees were also receiving reminders about the targets.

Methodologically, our work borrows from Augenblick et al. (2015). Their participants completed a series of real-effort tasks over three sessions spaced one week apart. In the week-1 session, participants decided how many tasks to allocate to the week-2 session and to the week-3 session. Then in the week-2 session, participants again allocated tasks between week 2 and week 3. In the week-2 session, any tasks allocated to week 2 had to be completed that same day, so present-biased participants were tempted to procrastinate and allocate fewer tasks to week 2. But in the week-1 session, because all work was in the future, the degree of a participant’s present-bias did not affect their allocation decision. The difference between the number of tasks allocated to week 2 in the week-1 session and the number of tasks allocated to week 2 in the week-2 session allowed the authors to estimate each individual’s present-bias parameter, $\beta$.

Conceptually, our paper is most closely related to Himmler et al. (2019)’s study on soft commitment devices in education. They implemented a system of planning and reminders to encourage MBA students at a German university to take the exams necessary for graduation. At the beginning of the semester, all students learned how to prepare for exams and which exams to take to stay on track for graduation. A control group received no additional information, one treatment group received two reminders of this information over the semester, and another treatment received reminders and signed a non-binding statement promising to follow the

---

6 One allocation decision was then selected probabilistically to be the “allocation-that-counts,” and participants had to complete the tasks according to the schedule outlined in the allocation-that-counts.

7 Additional details of the methodology in Augenblick et al. (2015) are discussed in Section 3.
recommended exam schedule. There was no additional penalty for failing to take the appropriate exams, so signing the planned exam schedule was a soft commitment device. Students in the commitment device treatment took and passed significantly more exams than both the control group and the reminders-only treatment group, and the effect was driven by individuals who were classified as procrastinators.

We apply the methods in Augenblick et al. (2015) to test the extent to which receiving a reminder of a plan reduces procrastination. Our work adds three main contributions to the literature. First, we make a methodological advance on Augenblick et al. (2015)’s study. In their experiment, participants were required to allocate tasks between two days, and only integer allocations of tasks were possible. For reasons discussed in Section 3.2, this feature contributed to violations of the weak axiom of revealed preference (Chakraborty et al., 2017). We address the problem by allowing participants to allocate fractional tasks between the two days. Second, while the existing research on planning and reminders focuses on smoothing effort over a long period of time, such as monthly workload or preparation for final exams, we study effort in a shorter time frame. Third, while the existing literature on planning has used externally imposed plans, our participants select their own optimal work schedule.

3) Experimental Methodology

We test the extent to which procrastination is reduced when an individual receives a reminder of their plan for when to complete a collection of real effort tasks. This question is an extension of Augenblick et al.’s study of procrastination, and the experimental design mirrors their methodology.

We conduct a three-day experiment, and participants must complete a series of tasks over Day 2 and Day 3 of the experiment. On Day 1, they decide how many tasks to allocate to
each day. Then on Day 2, they decide how many tasks to allocate to each day, again, and one allocation decision is selected at random to be the “allocation-that-counts.” In the treatment group, participants receive a reminder of their original allocation before making their second allocation decision on Day 2. We hypothesize that this reminder will reduce present-bias.

We use the difference between the number of tasks allocated to Day 2 on Day 1 and the number allocated to Day 2 on Day 2 to identify present-bias. If a participant allocated a greater share of tasks to Day 2 in the decision they made on Day 1 than in the decision they made on Day 2, then the participant was present-biased—that is, the participant procrastinated.

This section proceeds as follows: first, we describe the tasks that subjects complete; second, we describe the allocation decisions made by participants; third, we outline the treatment; and fourth, we explain the compensation structure and participant pool.

3.1) Tasks

Because our measurement of procrastination relies on the number of tasks a participant allocates to Day 2 when Day 2 is in the future compared to when Day 2 is the present, the measurement is more precise if the participants are required to allocate a greater number of tasks. To that end, subjects allocate 20-60 tasks. The exact number of tasks that a participant has to complete depends on their choices and on chance.

In order to make procrastination tempting, the tasks are tedious. As shown in Figure 1, subjects view a line of blurry Greek letters, and they must transcribe the letters with at least 80% accuracy onto a website. Each task takes approximately one minute to finish. Before participants allocate tasks between the two days, they have an opportunity to practice the task. This experience allows subjects to make an informed decision on how to allocate the tasks. After they complete the practice, they make a series of allocations between the two days.
3.2) Allocation decisions and task rates

On Day 1 and Day 2, participants make five allocation decisions at different “task rates.” A task rate, $R$, is the reduction in the number of tasks that must be completed on Day 2 in exchange for allocating an additional task to Day 3. For example, a task rate of 1:1.50 means that for every additional task allocated to Day 3, the tasks remaining for Day 2 falls by 1.50. Notice that if $R < 1$, then allocating one more task to Day 3 reduces the number of tasks left for Day 2 by less than one. Therefore, pushing off tasks to Day 3 is especially costly when the task rate is low, and participants who choose to do so have a high degree of impatience. The decision problem is to allocate tasks between Day 2 and Day 3 subject to the constraint

$$e_2 + Re_3 = 30$$

where $e_2$ and $e_3$ are the number of tasks allocated to Day 2 and 3, respectively. If the task rate $R$ is 1:1, then participants must allocate 30 tasks total, but notice that if $R=0.50$, a participant could choose to allocate 60 tasks to Day 3 and 0 to Day 2.

On Day 1, participants allocate the tasks at task rates 1:0.50, 1:0.75, 1:1.00, 1:1.25, and 1:1.50. On Day 2, participants again make allocation decisions at the same five task rates. One out of the ten allocations is chosen at random to be the allocation-that-counts, and a participant must complete the tasks according to the allocation-that-counts. Before making the first round of
allocation decisions on Day 1, subjects are told that they will have the opportunity to make a second round of decisions on Day 2, and that they will have to complete the tasks according to the schedule they outlined in the randomly selected allocation-that-counts. Allocations are made on a website using a sliding bar as displayed in Figure 2.

**Figure 2: Sample Allocation Decision**

For several task rates, fractional tasks are necessary to satisfy the constraint that $e_2 + Re_3 = 30$. Here we depart from the methods of Augenblick et al. (2015), where only integer allocations were possible. That restriction required rounding that ultimately led to violations of the weak axiom of revealed preference (Chakraborty et al., 2017). To allow for fractional tasks, we shorten a line of Greek text. A full-length task has 27 characters, so if a participant has to complete 0.33 tasks, they receive a line with nine characters.

---

8. Figure 2 refers to a “job rate” instead of a “task rate.” The word “task” has a specific meaning in the context of MTurk, so we use “job” instead of “task” on the experiment’s website.

9. For example, if fractional tasks were not available and the task rate was 1:1.50, then the allocations Day 2: 4 tasks / Day 3: 17 tasks and Day 2: 5 tasks / Day 3: 17 tasks would both be available. Subjects who allocated 5 tasks to Day 2 when 4 tasks were available violated WARP.
3.3) Treatment

We define the allocations participants made on Day 1 to be their plan for when to complete the tasks. Our hypothesis is that receiving a reminder of one’s plan reduces procrastination. On Day 2, to ensure that the participants in the treatment group focus on their plan, they are asked, “do you remember the allocations you made yesterday?” They fill out a table with estimates of their original allocation for all five task rates, and they earn $0.05 for every allocation they recall correctly. Then, they are reminded of their original plan by viewing the correct answers. Finally, they make the second round of allocation decisions.

The control group makes their allocation decisions on Day 2 before any discussion of their original allocations. After they complete the second round of allocations, they are asked to fill out the table with estimates of their original allocations, and they earn $0.05 for every line they get correct. Then, they are shown their original allocation.

3.4) Compensation structure and participants

After the participants make all ten allocations, one allocation is selected at random to be the allocation-that-counts. To make allocations incentive compatible, participants are compensated if they complete the tasks they allocated to Day 2 and Day 3 according to the allocation-that-counts.

To receive their payment, participants must also complete two “minimum work” tasks on Day 2 and Day 3 before they can begin the tasks that satisfy the allocation-that-counts. This minimum work requirement ensures that all subjects experience the fixed costs of logging on to the website and completing the first few tasks. If an individual allocates zero tasks to one of the days, he must nonetheless complete the two “minimum work” tasks.
Approximately two hundred participants are recruited from Amazon Mechanical Turk (MTurk) to complete the experiment, and they are randomly assigned between the treatment and control groups. Before they make any allocations, they must take a comprehension quiz on the instructions. All participants who complete the experiment and score 80% or above on the quiz earn at least $7.50, and participants who complete the experiment but score below 80% earn at least $6.00. If they drop out at any point after completing the quiz, they are paid $0.10.

4) Statistical Methodology

An individual’s present-bias is measured using the difference between the number of tasks allocated to Day 2 on Day 1 and the number of tasks allocated to Day 2 on Day 2. To test the extent to which receiving a reminder of one’s plan reduced procrastination, we calculate the average present-bias in the treatment and control groups. Each individual’s present-bias, or preference to procrastinate, is the value of $\beta$ in their intertemporal utility function. Following Augenblick et al., we estimate the average $\beta$ for each group. To do so, certain theoretical assumptions about the shape of subjects’ utility functions are necessary.

4.1) (Dis)utility function

The instantaneous cost of effort at time $t$ is of the form $(e_t + \omega)\gamma$, where $e_t$ is the number of tasks completed on day $t$. The parameter $\omega$ can be interpreted as the fixed costs associated with completing the first task. While $\omega$ does affect the curvature of the utility function, changes in $\omega$ have minimal effects on the value of $\beta$ (Andreoni & Sprenger, 2012). For simplicity, we set $\omega = 2$ to reflect the minimum work requirement of two tasks. The parameter $\gamma > 1$ ensures that

---

10 To assess the risk of attrition, a calibration group of 30 participants was used prior to the main study, and 1/3 of the calibration group completed the full experiment. To target 200 participants in the main study, the experiment was made available to six hundred workers on MTurk. Attrition rates are presented in Table 1.
the disutility function is convex such that first order conditions correspond to an optimal allocation, and the value of γ is calculated from the data. Accounting for hyperbolic discounting over time, the disutility of completing tasks over days 2-3 is

\[ D(e_2, e_3) = \beta \delta (e_2 + \omega)^\gamma + \beta \delta^2 (e_3 + \omega)^\gamma \] if \( t = 1 \)

\[ D(e_2, e_3) = (e_2 + \omega)^\gamma + \beta \delta (e_3 + \omega)^\gamma \] if \( t = 2 \)

If the current time \( t \) is Day 2, then Day 3 is the only future period, and therefore only Day 3 is discounted by \( \beta \).

An agent’s utility-maximization problems on Day 1 and Day 2 are

\[ \min_{e_2, e_3} \beta \delta (e_2 + \omega)^\gamma + \beta \delta^2 (e_3 + \omega)^\gamma \] if \( t = 1 \) \hspace{1cm} (1)

\[ \min_{e_2, e_3} (e_2 + \omega)^\gamma + \beta \delta (e_3 + \omega)^\gamma \] if \( t = 2 \) \hspace{1cm} (2)

subject to the constraint \( e_2 + R e_3 = 30 \). Notice that in equation (1), both terms in the disutility function are multiplied by \( \beta \) and \( \delta \). We could divide equation (1) by \( \beta \) and \( \delta \) and the solution to the minimization problem \((e_2^*, e_3^*)\) will be the same. Dividing out \( \beta \) and \( \delta \) allows from equation (1) allows equations (1) and (2) to be combined into a generalized utility-maximization problem:

\[ \min_{e_2, e_3} (e_2 + \omega)^\gamma + \beta^{1 \times 2} \delta (e_3 + \omega)^\gamma \]

Where \( 1_{t=2} \) is an indicator equal to 1 if \( t = 2 \). At this stage, we can rewrite the constraint as \( e_3 = \frac{1}{R} (30 - e_2) \) and substitute that expression into the minimization problem. Solving the minimization problem yields the first order condition for the optimal allocation \((e_2^*, e_3^*)\):

\[ \left( \frac{e_2^* + \omega}{e_3^* + \omega} \right)^{\gamma - 1} \frac{1}{\beta^{1 \times 2} \delta} = \frac{1}{R} \] \hspace{1cm} (3)

4.2) Identification
Taking logs of equation (3) and rearranging yields an equation that is linear in the observable variables: the log of the reciprocal of the task rate, log(1/R), and the indicator for whether allocation was made on Day 2, 1\text{t=2}.

\[
\log \left( \frac{e_2^* + \omega}{e_3^* + \omega} \right) = \frac{\log(\delta)}{\gamma-1} + \frac{\log(\beta)}{\gamma-1} (1_{t=2}) + \frac{1}{\gamma-1} \ast \log \left( \frac{1}{R} \right)
\]  

(4)

Because \omega=2 and the number of tasks allocated to Day 2 and Day 3 are observed, the dependent variable is observed. Thus, we can estimate the regression

\[
\log \left( \frac{e_2^* + \omega}{e_3^* + \omega} \right)_i = \eta_0 + \eta_1 (1_{t=2})_i + \eta_2 \log \left( \frac{1}{R} \right) + \epsilon_i
\]

where \eta_0 = \frac{\log(\delta)}{\gamma-1}, \eta_1 = \frac{\log(\beta)}{\gamma-1}, and \eta_2 = \frac{1}{\gamma-1}. Standard errors are clustered at the individual level.

We estimate this equation with a two-limit tobit regression because the dependent variable is censored from above and from below.\(^{11}\) Because \((e_2^*, e_3^*)\) is defined as the subject’s optimal allocation of tasks between the two dates, the dependent variable is determined by this optimal allocation. While we observe subjects’ selected allocation, if the selected allocation is at a corner solution (all tasks to Day 2 or all to Day 3), it is possible that the selected allocation does not satisfy the individual’s first-order conditions. In that case, the first order condition may not hold with equality, and the optimal allocation may be even more extreme.\(^{12}\)

The parameters of interest are recovered as \(\hat{\gamma} = 1 - \frac{1}{\hat{\eta}_2}\), \(\hat{\delta} = \exp \left( \frac{\hat{\eta}_0}{-\hat{\eta}_2} \right)\), and \(\hat{\beta} = \exp \left( \frac{\hat{\eta}_1}{-\hat{\eta}_2} \right)\). Because the parameters are non-linear combinations of the regression coefficients,


\(^{12}\) For example, if the observed allocation on Day 2 is 0 tasks on Day 2 and 30 on Day 3, then it is possible that the agent is extremely present-biased and to satisfy their first-order condition they would have to complete a negative quantity of tasks on Day 2 and more than 30 on Day 3. Because this allocation is not feasible, the value of the dependent variable is censored.
their standard errors are calculated using the delta method. Finally, we perform $\chi^2$ tests for the hypothesis that $\beta = 1$ and for the hypothesis that $\beta_{\text{treatment}} = \beta_{\text{control}}$.

5) Results

5.1) Summary Statistics: Participant Characteristics

Attrition rates were high in the initial stages of the experiment, but relatively low among participants who made a serious attempt to complete the experiment. While participants began the study on Amazon Mechanical Turk, the experiment itself was hosted on a separate website. To participate, subjects first submitted an informed consent form on MTurk, and then they were directed to the experiment’s website where they could create an account and complete Day 1 of the experiment. Many potential participants completed the MTurk informed consent form but did not create an account on the main website. As shown in Table 1, out of 616 workers who completed the informed consent form, 43.8% made an account on the experiment’s website and 37.2% completed Day 1. After completing the informed consent form, any additional compensation was earned only after completing the full experiment, so those who completed Day 1 were likely serious about completing all three days.

<table>
<thead>
<tr>
<th>Stage of Experiment</th>
<th>Frequency Completed</th>
<th>Percent Completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Informed Consent (MTurk)</td>
<td>616</td>
<td>100</td>
</tr>
<tr>
<td>Account Created</td>
<td>270</td>
<td>43.8</td>
</tr>
<tr>
<td>Day 1</td>
<td>229</td>
<td>37.2</td>
</tr>
<tr>
<td>Day 2 Allocations</td>
<td>212</td>
<td>34.4</td>
</tr>
</tbody>
</table>


14 The high attrition rates in the early stages of the experiment is unsurprising because workers earned $0.10 for completing the MTurk portion of the experiment, and this payment is relatively high for MTurk workers (Hara et al., 2018). Many potential participants likely took advantage of this deal and did not intend to complete the full experiment.
On Day 1, participants answered simple comprehension questions on the experiment’s instructions. Those who scored an 80% or above earned an additional $1.50 upon completing the experiment. Comprehension rates were high, with 91.5% of the control group and 92.8% of the treatment group scoring at least 80%.

On Day 1 and Day 2, participants’ allocation decisions revealed their preferences over different possible schedules for completing the tasks. Over half of our participants violated the Weak Axiom of Revealed Preference at least once. Recall, participants made allocation decisions at five different task rates. If $R = 0.50$, then allocating one more task to Day 3 reduces the number of tasks remaining for Day 2 by 0.50; allocating tasks to Day 2 is efficient. If $R = 0.75$, then allocating one more task to Day 3 reduces the number of tasks remaining for Day 2 by 0.75, so allocating tasks to Day 2 is more efficient when $R = 0.50$ than when $R = 0.75$. On each day that participants made allocations, the number of tasks allocated to Day 2 must be weakly increasing in $R$ to be consistent with WARP. Out of the 212 participants who completed both rounds of allocation decisions, 56.6% made at least one allocation that violated WARP.

<table>
<thead>
<tr>
<th>Table 2: WARP-Inconsistent Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
</tr>
<tr>
<td>WARP-Inconsistent</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>WARP-Consistent</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Notes: Frequency reported in parentheses.

5.2) Allocation Decisions and Parameter Estimates
We first examine the share of tasks a participant allocated to Day 2 on Day 1, and compare this value to the share of tasks they allocated to Day 2 on Day 2. Table 3 presents paired t-tests for the difference between these shares. For present-biased participants, the share of tasks allocated to Day 2 on Day 1 will be greater than the share allocated to Day 2 on Day 2, so the difference will be positive (as shown in columns (3), (6), and (9)). For time-consistent participants, the share of tasks allocated to Day 2 will be the same on Day 1 and on Day 2, so the difference will be zero.

Contrary to our hypothesis, subjects in the treatment group, who received a reminder of their original allocations before completing their second round of allocation decisions, showed evidence of procrastination significant at the 1% level, whereas subjects in the control group showed no evidence of procrastination. When the groups are aggregated, we find evidence of present bias significant at the 5% level for four out of the five task rates.

Table 3: Share of Tasks Allocated to Day 2
Panel A: Full Sample

<table>
<thead>
<tr>
<th>R</th>
<th>Control</th>
<th></th>
<th></th>
<th>Treatment</th>
<th></th>
<th></th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Choice on Day 1</td>
<td>Choice on Day 2</td>
<td>Dif</td>
<td>Choice on Day 1</td>
<td>Choice on Day 2</td>
<td>Dif</td>
<td>Choice on Day 1</td>
</tr>
<tr>
<td>0.50</td>
<td>.532</td>
<td>.529</td>
<td>.003</td>
<td>.573</td>
<td>.509</td>
<td>.065***</td>
<td>.552</td>
</tr>
<tr>
<td></td>
<td>(.287)</td>
<td>(.314)</td>
<td>(.026)</td>
<td>(.250)</td>
<td>(.300)</td>
<td>(.021)</td>
<td>(.269)</td>
</tr>
<tr>
<td>0.75</td>
<td>.555</td>
<td>.569</td>
<td>-.013</td>
<td>.583</td>
<td>.516</td>
<td>.067***</td>
<td>.569</td>
</tr>
<tr>
<td></td>
<td>(.274)</td>
<td>(.298)</td>
<td>(.026)</td>
<td>(.234)</td>
<td>(.292)</td>
<td>(.024)</td>
<td>(.255)</td>
</tr>
<tr>
<td>1.00</td>
<td>.507</td>
<td>.498</td>
<td>.008</td>
<td>.540</td>
<td>.454</td>
<td>.087***</td>
<td>.523</td>
</tr>
<tr>
<td></td>
<td>(.235)</td>
<td>(.256)</td>
<td>(.025)</td>
<td>(.191)</td>
<td>(.227)</td>
<td>(.023)</td>
<td>(.214)</td>
</tr>
<tr>
<td>1.25</td>
<td>.443</td>
<td>.414</td>
<td>.029</td>
<td>.465</td>
<td>.391</td>
<td>.075***</td>
<td>.454</td>
</tr>
<tr>
<td></td>
<td>(.277)</td>
<td>(.281)</td>
<td>(.024)</td>
<td>(.254)</td>
<td>(.263)</td>
<td>(.024)</td>
<td>(.265)</td>
</tr>
<tr>
<td>1.50</td>
<td>.448</td>
<td>.430</td>
<td>.017</td>
<td>.486</td>
<td>.401</td>
<td>.085***</td>
<td>.467</td>
</tr>
<tr>
<td></td>
<td>(.277)</td>
<td>(.297)</td>
<td>(.029)</td>
<td>(.270)</td>
<td>(.272)</td>
<td>(.028)</td>
<td>(.274)</td>
</tr>
</tbody>
</table>

Obs 106 106 106 106 106 106 212 212 212

Panel B: Restricted Sample: WARP-Consistent
### Panel C: Restricted Sample: WARP-Consistent, High Comprehension, and Completed Experiment

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>.608</td>
<td>.597</td>
<td>.012</td>
<td>.641</td>
<td>.596</td>
<td>.046**</td>
<td>.626</td>
<td>.597</td>
<td>.030</td>
</tr>
<tr>
<td></td>
<td>(.289)</td>
<td>(.311)</td>
<td>(.032)</td>
<td>(.245)</td>
<td>(.280)</td>
<td>(.022)</td>
<td>(.266)</td>
<td>(.293)</td>
<td>(.019)</td>
</tr>
<tr>
<td>0.75</td>
<td>.62</td>
<td>.609</td>
<td>.011</td>
<td>.635</td>
<td>.592</td>
<td>.044*</td>
<td>.628</td>
<td>.6</td>
<td>.028</td>
</tr>
<tr>
<td></td>
<td>(.270)</td>
<td>(.303)</td>
<td>(.030)</td>
<td>(.234)</td>
<td>(.268)</td>
<td>(.023)</td>
<td>(.251)</td>
<td>(.284)</td>
<td>(.018)</td>
</tr>
<tr>
<td>1.00</td>
<td>.549</td>
<td>.519</td>
<td>.029</td>
<td>.515</td>
<td>.489</td>
<td>.027</td>
<td>.531</td>
<td>.504</td>
<td>.028</td>
</tr>
<tr>
<td></td>
<td>(.220)</td>
<td>(.241)</td>
<td>(.026)</td>
<td>(.177)</td>
<td>(.188)</td>
<td>(.022)</td>
<td>(.198)</td>
<td>(.214)</td>
<td>(.018)</td>
</tr>
<tr>
<td>1.25</td>
<td>.418</td>
<td>.374</td>
<td>.044*</td>
<td>.424</td>
<td>.395</td>
<td>.03</td>
<td>.421</td>
<td>.385</td>
<td>.037**</td>
</tr>
<tr>
<td></td>
<td>(.293)</td>
<td>(.292)</td>
<td>(.026)</td>
<td>(.250)</td>
<td>(.258)</td>
<td>(.022)</td>
<td>(.270)</td>
<td>(.273)</td>
<td>(.017)</td>
</tr>
<tr>
<td>1.50</td>
<td>.411</td>
<td>.360</td>
<td>.051*</td>
<td>.415</td>
<td>.394</td>
<td>.022</td>
<td>.413</td>
<td>.378</td>
<td>.036*</td>
</tr>
<tr>
<td></td>
<td>(.302)</td>
<td>(.292)</td>
<td>(.028)</td>
<td>(.257)</td>
<td>(.261)</td>
<td>(.025)</td>
<td>(.278)</td>
<td>(.276)</td>
<td>(.018)</td>
</tr>
</tbody>
</table>

| Obs | 56 | 56 | 56 | 64 | 64 | 64 | 120 | 120 | 120 |

Notes: * p < 0.10, ** p < 0.05, *** p < 0.01. Stars are reported only for columns (3), (6), and (9). Standard errors reported in parentheses.

Given that over half of the participants made at least one allocation decision that violated WARP, we may also be interested in the results when we limit the sample to participants who made only WARP-consistent decisions. Those who violated WARP are more likely to have misunderstood the trade-off implied by the task rates. If they had fully understood the task rates, they may have made different allocation decisions, and therefore their allocation decisions do not
necessarily reflect their true preferences. Panel B repeats the analyses for the sample of WARP-consistent participants.

After restricting to only WARP-consistent participants, Panel B tells a different story. While the treatment group showed highly significant evidence of procrastination at all five task rates in Panel A, in Panel B, the treatment group procrastinated significantly at only two task rates (0.50 and 0.75) and all p-values increased substantially. This decline in significance is not due to an increase in the standard errors; standard errors remained nearly constant between Panels A and B despite the reduction in sample size. Instead, the results lose much of their significance because the value of interest, the difference between the share of tasks allocated to Day 2 on Day 1 and the share allocated to Day 2 on Day 2, falls by over 50%, on average. Comparing Panels A and B for the control group, we observe that while the control group showed no evidence of procrastination in Panel A, the subsample of WARP-consistent participants did show some evidence of procrastination for two of the task rates (1.25 and 1.50), with p < 0.10. In Panel B, we have limited evidence of procrastination in the control group for task rates 1.25 and 1.50, and limited evidence of procrastination in the treatment group for task rates 0.50 and 0.75. Taken together, these results suggest no effect of the reminder on procrastination.

Finally, we examine one more subsample. The primary advantage of restricting to participants who made WARP-consistent decisions was that those participants were more likely to have understood the nature of the decision problem they faced, and therefore their allocation decisions were more likely to reflect their true preferences. We can also take advantage of additional criteria to exclude participants who may not have understood the experiment. We re-estimate our results for participants who were WARP-consistent, scored at least 80% on the
comprehension quiz on Day 1, and completed all three days of the experiment. The comprehension quiz directly tests a participant’s understanding of the instructions, so including that criterion is a natural way to exclude those who did not adequately understand the experiment. This restriction removes four participants from the previous sample, two each from the treatment and control. Excluding participants who failed to complete all three days of the experiment is also likely to exclude those who did not adequately understand the experiment. If a participant made allocation decisions but did not complete the tasks, then they may not have understood the difficulty of the tasks, or they may not have understood that their allocation decisions were going to determine the number of tasks they would have to complete. They might have made different allocation decisions if they had better understood the requirements of the experiment. This criterion excludes only two additional participants, one each from the treatment and control. Panel C of Table 3 presents these results, and Panel C closely resembles Panel B. The main difference is that the procrastination in the control group loses its significance. We explore the reason for this change below when we discuss the parameter estimates in Panel C of Table 4.

Two additional trends in Table 3 are worth noting. First, recall that as the task rate increases, allocating tasks to Day 2 becomes less efficient. So, as the task rate increases, we should expect the share of tasks allocated to Day 2 to decrease. But in all three panels, while on average the share of tasks allocated to Day 2 is decreasing in $R$, it is not uniformly decreasing. When we restrict the sample to WARP-consistent participants in Panels B and C, although the share allocated to Day 2 is still not uniformly decreasing in $R$, we do find a steeper profile of the share allocated to Day 2 as $R$ increases. The second notable feature of Table 3 is that participants
are fairly patient. When the task rate was 1:1, all subsamples allocated at least 50% of the tasks to Day 2 in the decision made on Day 1.

Table 4 presents the parameter estimates from our identification strategy. The estimates of $\beta$ mirror the findings from Table 3. Considering the full sample, Panel A of Table 4 rejects the null hypothesis that $\beta = 1$ for the treatment group with $p < 0.01$, consistent with the finding in Table 3 that the full treatment group procrastinated significantly. And for the control group, Panel A of Table 4 shows no evidence of present-bias, consistent with Panel A of Table 3. Table 4 also indicates that we cannot reject the null hypothesis that $\beta_{\text{treatment}} = \beta_{\text{control}}$. That is, the reminder did not significantly reduce present-bias.

### Table 4: Parameter Estimates

#### Panel A: Full Sample

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Treatment</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>.879</td>
<td>.561</td>
<td>.703</td>
</tr>
<tr>
<td></td>
<td>(.162)</td>
<td>(.160)</td>
<td>(.109)</td>
</tr>
<tr>
<td>$\delta$</td>
<td>.917</td>
<td>1.127</td>
<td>1.015</td>
</tr>
<tr>
<td></td>
<td>(.153)</td>
<td>(.134)</td>
<td>(.098)</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>1.817</td>
<td>1.783</td>
<td>1.800</td>
</tr>
<tr>
<td></td>
<td>(.276)</td>
<td>(0.284)</td>
<td>(.198)</td>
</tr>
</tbody>
</table>

$H_0$: $\beta = 1$

<table>
<thead>
<tr>
<th></th>
<th>$\chi^2$</th>
<th>$p$</th>
<th>$\chi^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.560</td>
<td>.4550</td>
<td>7.570</td>
<td>.0059</td>
</tr>
<tr>
<td>$\beta$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$H_0$: $\beta_{\text{treatment}} = \beta_{\text{control}}$

<table>
<thead>
<tr>
<th></th>
<th>$\chi^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.96</td>
<td>.161</td>
</tr>
</tbody>
</table>

| Observations | 1060 | 1060 | 2120 |
| Clusters     | 106  | 106  | 212  |

#### Panel B: Restricted Sample: WARP-Consistent

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Treatment</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.856</td>
<td>.876</td>
<td>.868</td>
</tr>
<tr>
<td></td>
<td>(.116)</td>
<td>(.097)</td>
<td>(.074)</td>
</tr>
<tr>
<td>$\delta$</td>
<td>1.047</td>
<td>1.032</td>
<td>1.038</td>
</tr>
<tr>
<td></td>
<td>(.132)</td>
<td>(.081)</td>
<td>(.074)</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>2.647</td>
<td>2.523</td>
<td>2.579</td>
</tr>
<tr>
<td></td>
<td>(0.506)</td>
<td>(.434)</td>
<td>(.329)</td>
</tr>
</tbody>
</table>

$H_0$: $\beta = 1$

<table>
<thead>
<tr>
<th></th>
<th>$\chi^2$</th>
<th>$p$</th>
<th>$\chi^2$</th>
<th>$p$</th>
<th>$\chi^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.540</td>
<td>.4550</td>
<td>1.650</td>
<td>.199</td>
<td>3.150</td>
<td>.076</td>
</tr>
</tbody>
</table>

| Observations | 2120 | 2120 | 2120 |
| Clusters     | 212  | 212  | 212  |
H₀: \( \beta_{\text{treatment}} = \beta_{\text{control}} \) \( \chi^2 = 1.96 \) \( p = .215 \)

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Treatment</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta )</td>
<td>.928</td>
<td>.876</td>
<td>.901</td>
</tr>
<tr>
<td></td>
<td>(.085)</td>
<td>(.097)</td>
<td>(.065)</td>
</tr>
<tr>
<td>( \delta )</td>
<td>.980</td>
<td>1.032</td>
<td>1.006</td>
</tr>
<tr>
<td></td>
<td>(.112)</td>
<td>(.081)</td>
<td>(.067)</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>2.728</td>
<td>2.608</td>
<td>2.663</td>
</tr>
<tr>
<td></td>
<td>(.531)</td>
<td>(.464)</td>
<td>(.348)</td>
</tr>
</tbody>
</table>

\( H_0: \beta = 1 \) \( \chi^2 = 0.720 \) \( p = 0.397 \) \( \chi^2 = 2.290 \) \( p = 0.130 \)

Notes: Standard errors in parentheses.

In Panel B of Table 4, we restrict to WARP-consistent participants, and the values of \( \beta \) in the treatment and control are nearly identical. We fail to reject the hypothesis that the degree of present-bias in the treatment group equals the degree of present-bias in the control. When testing each value of \( \beta \) separately for the treatment and control groups, neither value is significantly smaller than one. However, when the treatment and control groups are aggregated, we can reject the hypothesis that \( \beta = 1 \) at the 10% level (\( p=0.076 \)). In this sample, we tentatively replicate the literature’s finding that individuals have present-biased preferences with respect to effort.

In Panel C of Table 4, we restricted to individuals who were WARP-consistent, had a high score on the comprehension quiz, and completed the experiment. In this sample, the estimate of \( \beta \) in the treatment group was unchanged from Panel B, but the value of \( \beta \) for the control group increased substantially from 0.856 in Panel B to 0.928 in Panel C. This change is driven by one participant in the control group who displayed the maximum amount of present
bias. In Tables 3 and 4, three participants are excluded from the control group when moving from Panel B to Panel C. Excluding two of those participants has little effect on the estimate of $\beta$, but excluding the individual who was maximally present-biased causes $\beta$ to increase to 0.928. This outcome suggests that our findings are sensitive to highly present-biased participants. Under this sample, we cannot reject the null hypothesis that $\beta = 1$ even after aggregating the treatment and control groups.

6) Discussion

There were two main findings that we were interested in, but we were unable to find conclusive evidence of either result. First, despite largely following the methods of Augenblick et al. (2015), we did not clearly replicate their key finding that individuals have present-biased preferences over real effort tasks. Second, we found no evidence that reminders alleviated procrastination problems, unlike to Himmler et al. (2019) and Cadena et al. (2011). Sections 6.1 and 6.2 explore several possible explanations for these differences.

6.1) Comparison with Augenblick et al. (2015)

We consider three potential reasons that we did not unequivocally replicate the finding that individuals are present-biased when completing real effort tasks. First, our participants may not have understood their allocation decisions as well as the participants in the original experiment, or they may have been inattentive. A greater level of confusion or inattention would introduce more noise and potentially prevent us from rejecting the hypothesis that $\beta = 1$. One source of confusion or inattention for our participants may have been that our experiment was

---

15 One explanation is that our participants simply were not present-biased. However, in a survey of 207 MTurk workers, Goodman et al. (2013) find that MTurk participants were significantly present-biased.
conducted online and asynchronously, while the original experiment was conducted in person. Conveying the instructions may have been more challenging in writing without a researcher present to read the instructions out loud, and the online format did not easily allow participants to ask questions if they needed clarification. Another source of inattention or confusion may be that our participants were workers on MTurk, while participants in the original experiment were undergraduate students at an UC Berkeley. Perhaps the allocation decisions were too complex for some of our participants to understand, but not too complex for the participants in Augenblick et al. (2015). The participants in the original study may have been more motivated to understand the decision problem that they were facing. If the experiment was run by a professor at their university, participants may have wanted to demonstrate to the professor that they were a smart group of students, whereas that pressure is much less likely to apply to MTurk workers. The data on WARP-inconsistent individuals is consistent with this “confusion/inattention” explanation. In our sample, 11.5% of allocation decisions were WARP-inconsistent and 56.6% of participants violated WARP at least once, but those numbers fall to 7.88% and 31.5%, respectively, for participants in the original experiment. While these differences are modest, they are the direction predicted by a confusion/inattention explanation.

A second possible reason that we did not conclusively replicate the finding that individuals are present-biased is that we did not adequately separate work in the present from work in the future. One concern is that the one-day horizon may not have been not far enough in the future for the work to be discounted by β; participants may not have thought of the tasks allocated to Day 2 on Day 1 as work in the future. In the original study by Augenblick et al. (2015), Day 1 and Day 2 were one week apart, so this concern was less serious. Alternatively, participants may have felt as though tasks allocated to Day 2 on Day 2 were not work in the
present because participants could make allocation decisions on Day 2 and come back hours later to complete the tasks. If a participant perceived work several hours in the future as work in a non-present period, then they would not have been tempted to procrastinate on Day 2.

Finally, a third possibility is that the reminder treatment was highly effective at reducing present bias, and the control group was also treated. On Day 2, the control group first made their allocation decisions, and then viewed a reminder of their original allocations. But due to technical limitations, the allocation decisions and the reminder were presented on the same webpage, and members of the control group could view the reminder if they scrolled far enough down the page. If the control group focused on the reminder before they submitted their allocations and the reminder was effective at reducing procrastination, then our participants would not appear to have present-biased preferences.

6.2) Comparison with commitment and reminders literature

The data did not provide evidence for our hypothesis that receiving a reminder would reduce present bias. We discuss three possible reasons that we could find this data even if the hypothesis is correct.16

First, as discussed in the previous section, if the control group viewed the reminder before making their second round of allocations, the control group would have been treated. In this case, the reminder may have been highly effective at reducing procrastination, but we would not be able to pick up an effect.

Second, the reminder may not have functioned as a commitment device. Participants may have felt little or no guilt for failing to follow their original plan, and the psychological penalty to

---

16 Of course, the hypothesis that reminders reduce present-bias may be false, or participants may not be present-biased even before receiving reminders.
procrastinating was too small to influence behavior. Alternatively, participants may not have conceptualized the allocations they made on Day 1 as a plan for when they ought to complete the tasks, in which case they experienced no guilt from making different allocation decisions on Day 2.

A third explanation is that the stakes in this experiment were too low. In Himmler et al. (2019) and Cadena et al. (2011), interventions aimed to reduce procrastination in preparing for exams and in monthly workload, respectively, and these interventions targeted major aspects of individuals’ lives. In the low-stakes environment of our three-day experiment, smoothing workload may be less important to participants and therefore a commitment device may be less effective.\textsuperscript{17}

7) Conclusion

We conducted a three-day experiment to test the extent to which a system of planning and reminders could function as a commitment device and reduce present-bias. In contrast to existing literature (Cadena et al. 2011; Himmler et al. 2019), we found no evidence that participants who received a reminder of their planned work schedule were less present-biased than those who received no reminder. In addition, although our estimates of $\beta$ were consistently less than one, we were not able to clearly replicate the finding that individuals are present-biased with respect to real effort tasks (Augenblick et al., 2015). One can take these findings at face value—perhaps our participants were not present-biased and/or the reminder had no effect on their behavior. Our inability to reject the hypothesis that $\beta = 1$ should be interpreted with

\textsuperscript{17} However, the effect could also go in the opposite direction. Individuals may be careful to avoid procrastinating in high-stakes situations even without a commitment device, so introducing a commitment device could have smaller effects when the stakes are higher.
skepticism, however, especially given the sensitivity of our results to highly present-biased individuals and the widespread documentation of present-bias in many contexts (Bryan et al., 2010). Regarding our main hypothesis that reminders will reduce present-bias, we find the lack of evidence supporting that conclusion to be more credible. It is quite plausible that receiving reminders about one’s plan was not an effective commitment device in our experiment because it did not generate a meaningful psychological penalty to procrastinating, or because the stakes of the experiment were too low. The conditions under which a system of planning and reminders reduce procrastination remains an area for future research.


Appendix

A1) Instructions on Day 1
All instructions are modified from Augenblick et al. (2015).

Introduction (1/6)

Eligibility for this study
Thank you for participating in our experiment. To be in this study, you need to meet these criteria:

- You must be willing to participate for three consecutive days.
- Participation will require your signing onto the experiment’s website for three consecutive days for at least 2 minutes per day and for an average of 60 minutes all together.
- You must be willing to receive your payment from this experiment as one single completion payment at the end of the study. Payments will be made within 24 hours after the final session, two days from today. If you do not meet these criteria, please exit this experiment.

Anonymity
You will receive reminders by email over the next two days. Your email information will not be recorded or connected to your responses in the experiment.

Your Earnings
Most of your earnings are based on two things: your completion of the experiment and your performance on a comprehension quiz about these instructions. If you complete all three days of the experiment and score an 80% or above on the comprehension quiz, you will receive $7.50. If you complete all three days of the experiment and score below 80% on the quiz, then you will receive $6.00. If you choose to end your participation before the three days are complete, your earnings will be $0.10. You can also earn up to $0.25 from an additional small task.

Jobs (2/6)
In this study you will be asked to complete a series of jobs. These jobs will be completed over time. Some portion of the jobs may be completed sooner, and some portion of the jobs may be completed later depending on your choices and chance. Importantly, some jobs must be completed every day. That is, as mentioned before, your participation is required on all three days of the study.

Transcribing greek letters
Each job asks you to transcribe letters from a line of greek text. Greek text will appear in the box on your screen. For each letter you will need to click on the corresponding letter to enter it onto the line on your screen. One job is one line of greek text. For the job to be complete your accuracy must be 80% or better.

Practice
Please spend a few minutes practicing the job. You will need to complete the practice before moving on.
**Experimental Timeline (3/6)**

Now that you’ve practiced the job, let’s consider the timeline of the study. Along the way we will discuss a few important details of how the study works.

Note: Minimum Work for each week: On Day 2 and Day 3, you will be required to complete a minimum number of jobs. This “minimum work” will be 2 jobs every day.

**Day 1 (today)**

Today, you will be asked to make a series of 5 decisions. In these decisions you are asked to allocate jobs between tomorrow (Day 2) and two days from today (Day 3). In each decision you are free to allocate your jobs as you choose. Note that this allocation decision does not include the minimum work for each day, which you must also complete. You will choose by moving a slider to your desired allocation. Please take a moment to practice.

**Job Rates**

In the example decision above every job you allocate to Day 3 reduces the number of jobs allocated to Day 2 by one. This is what we will refer to as a 1:1 job rate. The job rate will vary across your decisions. For example, the job rate may be 1:1.50, such that every job you allocate to Day 3 reduces the number of jobs allocated to Day 2 by 1.50. Or, the job rate may be 1:0.50, such that every job you allocate to Day 3 reduces the number of jobs allocated to Day 2 by 0.50. For simplicity, the job rates will always be represented as 1:X, and you will be fully informed of the value of X when making your decisions. Please practice with the different job rates now.

Slider setup: To set up the sliders, first drag the dot to the left end and then bring it back to the right. If you did the setup correctly, you should see Day 2: 30.0 and Day 3: 0.0 when the dot is all the way to the right.

You may notice that some allocations have a fractional number of jobs. For example, if 0.66 jobs are allocated to Day 3, then you will only have to transcribe 66% of a line of text.

**Day 2 (tomorrow)**

On Day 2, one day from today, you will receive an email with instructions on how to access the website. You will again complete your minimum work of 2 jobs. You will be asked again to make 5 allocation decisions. Exactly one of your 10 total allocation decisions will be
implemented. That is, we will implement one decision from Day 1 or one decision from Day 2. Each decision has an equal chance of being selected. We refer to the allocation decision that gets selected as the “allocation-that-counts.” The jobs you allocated to Day 2 in the allocation-that-counts must be completed on Day 2. If you do not return or do not complete the jobs on Day 2, you cannot complete the study, and you will receive only the minimum payment of $0.10. In order for your jobs on Day 2 to be counted, they must be submitted by 11:59pm tomorrow.

Day 3 (day after tomorrow)
Day 3, two days from today, will also occur on the experiment’s website, and you will receive an email with instructions on how to access the website. You will again complete your minimum work of 2 jobs. Then, you must complete the jobs you allocated to Day 3 in the allocation-that-counts. If you do not return or do not complete the jobs on Day 3, you cannot complete the study, and you will receive only the minimum payment of $0.10. In order for your jobs on Day 3 to be counted, they must be submitted by 11:59pm two days from now.

REMEMBER: EACH ALLOCATION DECISION COULD BE THE ALLOCATION-TThat- COUNTS SO TREAT EACH DECISION AS IF IT WAS THE ONE DETERMINING YOUR JOBS.

Recap (4/6)
- You will be participating in a study that requires participation each day for three consecutive days, starting today.
- You will complete a series of jobs, and each job requires you to transcribe greek letters.
- On all three days, you will be asked to complete the minimum work of 2 jobs.
- On Day 1, today, you will be asked to make a series of allocation decisions. You will allocate jobs to Days 2 and 3 at various job rates.
- On Day 2, you will again make allocation decisions.
- One of your 10 allocation decisions will be chosen at random to be the allocation-that-counts and that allocation will determine the jobs that you complete on Days 2 and 3.
- If you score 80% or higher on the following quiz and complete the three-day experiment, you will receive a payment of at least $7.50 after the end of Day 3.
- If you score lower than 80% on the following quiz and complete the three-day experiment, you will receive a payment of at least $6.00 after the end of Day 3.
- If you choose to no longer participate, or do not complete the jobs you chose, you will receive $0.10.

Quiz (5/6)
Please complete the quiz in order to make sure that you understand the allocation decisions and the timeline of the study. If you score an 80% or above and complete the rest of the experiment, you will earn at least $7.50. If you score below 80% and complete the rest of the experiment, you will earn at least $6.00.

How many days are you required to participate? (include today)
- 1
- 2
- 3
How many jobs do you have to complete each day to satisfy “minimum work”?

- 1
- 2
- 3
- 4

If you face a 1:2 job rate for allocations between Days 2 and 3, every job you allocate to Day 3 decreases by how many the number of jobs you allocate to Day 2?

- 1
- 1.5
- 2
- 2.5

The “allocation that counts” will be one allocation decision out of ___ possibilities:

- 3
- 10
- 15
- 50

What are the jobs you will have to complete on Days 2 and 3?

- Slide a dot along a bar
- Count the number of letters in a sentence
- Alphabetize a list of words
- Transcribe blurry greek letters

Allocations (6/6)

Reminder of Timeline
Today you will be asked to make a series of 5 allocation decisions for the transcription jobs. In these decisions you are asked to allocate jobs between tomorrow (Day 2) and two days from today (Day 3). In each decision you are free to allocate your jobs as you choose. The allocations do not include the “minimum work” requirement. You will choose by moving a slider to your desired allocation.

Making allocations
In the sliders on the screen, you will be asked to make 5 allocations. Remember each allocation decision could be the allocation-that-counts, so please make each decision as if it were the one that determines your jobs.
A2) Instructions on Day 2
The instructions for the treatment group are presented below.

Overview

Day 2
Welcome back to our experiment. Today involves four steps:
- You will complete the minimum work requirement of two jobs.
- You will complete a memory test.
- You will make 5 allocation decisions of jobs between Day 2 (today) and Day 3 (tomorrow).
- We will select the “allocation-that-counts” and you will complete the number of jobs that were allocated to Day 2 according to the allocation-that-counts.

Minimum work
Now you will complete your minimum work of two jobs.

Extra task: short memory quiz
Do you remember the allocation decisions you made yesterday? Please fill out the following table with your best estimates. You will earn $0.05 for every row you get correct.

---

18 For the control group, this stage comes after they submit their allocation decisions on Day 2.
19 Participants selected a value from the dropdown in one cell, and the other cell in that row was filled in automatically.
When finished, complete the remaining work.²⁰

At this point, we have selected the allocation-that-counts. Your allocation-that-counts is the decision you made on Day X with job rate Y. You must complete A jobs today and B jobs tomorrow to receive the rest of your compensation. Please begin today’s jobs.

A3) Instructions on Day 2

Overview

²⁰ Participants then click on a button that takes them to the next page.
Day 3
Welcome back to our experiment. Today involves two steps:
- You will complete the minimum work requirement of two jobs.
- Complete the jobs allocated to Day 3.

**Minimum Work**
Please complete your minimum work of two jobs.

**Remaining Work**
Now please complete the remaining B jobs for Day 3. Each job should take around 1 minute. If you submit these jobs by 11:59pm, you will receive your compensation of $6 plus up to $1.75 in earned bonuses.

**When all jobs are completed, click submit.**

This experiment was designed to test the effect of reminders on procrastination. You can think of the allocations you made on Day 1 as your original plan for when to complete the jobs, and we measured procrastination using the difference between the allocations you made on Day 1 and the allocations you made on Day 2. The memory quiz on Day 2 was a reminder of your original plan. One group of workers saw the memory quiz (reminder) before they made their allocations on Day 2, and the other group saw the memory quiz after they made their allocations on Day 2. By comparing the two groups, we can identify the effect of the reminder on procrastination. Procrastinating more or procrastinating less did not affect anyone's earnings.

If you have any questions or concerns about this experiment, you can contact the Haverford Institutional Review Board at hc-irb@haverford.edu

---

21 Hitting submit takes participants to a new page.